Supporting Material for "MCMC can detect non-identifiable models"

Ivo Siekmann[1]
Auckland Bioengineering Institute,
The University of Auckland, Auckland, New Zealand

James Sneyd
Department of Mathematics,
The University of Auckland, Auckland, New Zealand

Edmund J. Crampin
Auckland Bioengineering Institute & Department of Engineering Science,
The University of Auckland, Auckland, New Zealand

[1]Corresponding author. Address: Auckland Bioengineering Institute, The University of Auckland, 70 Symonds St, Auckland, New Zealand, Tel.: (+64) 9 373 7599 ext. 81863, Fax: (+64)9 367 7157

1 Fitting models with one open or one closed state

If the model has only one open state (but possibly several closed states) the efficiency of the algorithm based upon the likelihood given by Eq. 14 in the main text can be considerably increased. An algorithm for this situation (which can be seen as a special case of the method presented here) was developed by Gin et al. (1, section 2.5). Of course, models with one closed (but several open states) can be treated analogously.

The crucial idea is that the complete sequence E of observations can be broken down into shorter subsequences that start and end in an open observation. Indeed, whenever an open event is observed, it is clear that the model is in its single open state and the preceding events have no influence on the rest of the sequence due to the Markov property. From this it is easy to see that the likelihood of a sequence of several consecutive open observations (O^n) is the same as the probability of observing $n-1$ open-open transitions:

$$\mathbb{P}((O^{n-1})|O,Q) = p_{OO}^{n-1} \tag{1}$$

where p_{OO} is the transition probability from the open to the open state. Since all subsequences are independent, the likelihood $\mathbb{P}_{\text{open}}$ for all open events observed in the sequence E can be calculated from the total number of open-to-open transitions N_{OO}:

$$\mathbb{P}_{\text{open}} = p_{OO}^{N_{OO}}. \tag{2}$$

The likelihood $\mathbb{P}(OC^n|O,Q)$ for all subsequences of the form (OC^nO) where C^n stands for n consecutive closed events can be obtained from

$$\mathbb{P}((OC^n)|O,Q) = (0,\ldots,1)\cdot(A_\tau\cdot P_C)^n\cdot A_\tau\cdot P_O\cdot u. \tag{3}$$

The probability $\mathbb{P}_{\text{closed}}$ then is

$$\mathbb{P}_{\text{closed}} = \prod_{i=1}^{M}\mathbb{P}((OC^i)|O)^{N(OC^iO)} \tag{4}$$

where $N(OC^iO)$ is the number of subsequences (OC^iO) containing i closed observations. Note that the number of matrix multiplications that are required is reduced to the maximum length M of consecutive closed observations as opposed to the total number N if calculated by Eq. 14 in the main text.

References

1. Gin, E., M. Falcke, L. E. Wagner, D. I. Yule, and J. Sneyd, 2009. Markov chain Monte Carlo fitting of single-channel data from inositol trisphosphate receptors. *Journal of Theoretical Biology* 257:460–474.